Datos circulares con R

María Alonso Pena

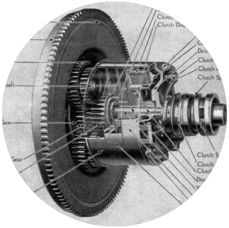
Departamento de Estatística, Análise Matemática e Optimización

Universidade de Santiago de Compostela

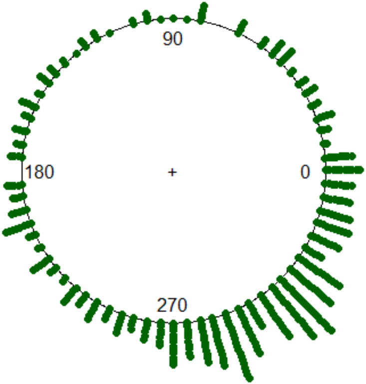
Datos circulares



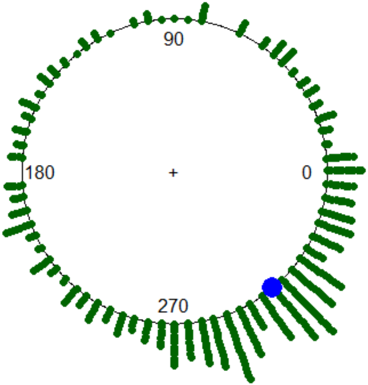




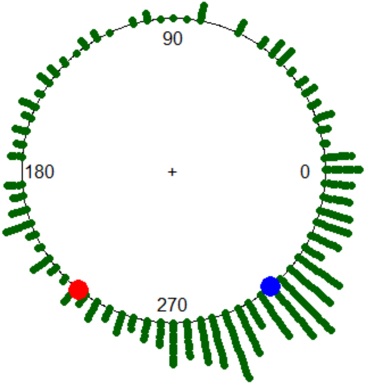
Pulgas de praia



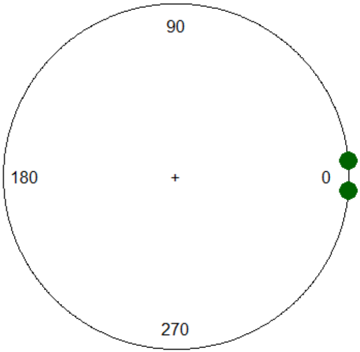
Pulgas de praia



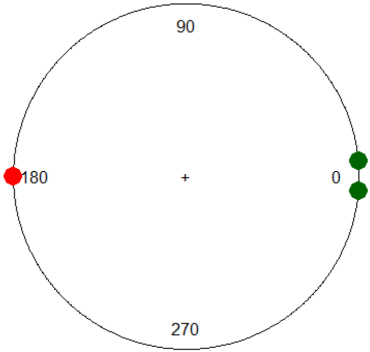
Pulgas de praia



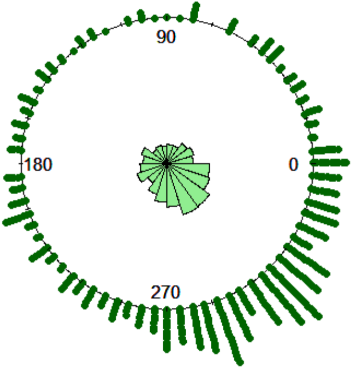
Un exemplo sinxelo



Un exemplo sinxelo



Representando datos circulares

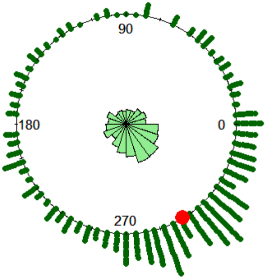



```
{circular}
pulgas<-circular(pulgas,units="degrees")
plot(pulgas,stack=T,col="darkgreen")
rose.diag(pulgas,add=T, col="lightgreen")
```

A media circular

$$\bar{\mu} = \mathsf{atan2}\left(\frac{1}{n}\sum_{i=1}^{n}\sin\theta_i, \frac{1}{n}\sum_{i=1}^{n}\cos\theta_i\right)$$

$$\bar{\mu} = 301.0753$$

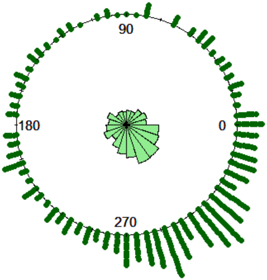



```
{circular}
mean.circular(pulgas)
```

Concentración dos datos

$$\bar{R} = \left(\left(\frac{1}{n} \sum_{i=1}^{n} \sin \theta_i \right)^2 , \left(\frac{1}{n} \sum_{i=1}^{n} \cos \theta_i \right)^2 \right)^{1/2}$$

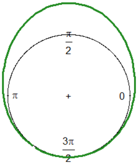
$$\bar{R} = 0.4433$$




```
{circular}
rho.circular(pulgas)
```

Densidade circular: von Mises

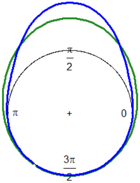
$$f(\theta;\mu,\kappa) = \frac{1}{2\pi I_0(\kappa)} \exp(\kappa \cos(\theta - \mu))$$




```
{circular}
mu<-circular(pi/2)
kappa<-2
curve.circular(dvonmises(x,mu,kappa),lwd=3,col="forestgreen")
```

Densidade circular: von Mises

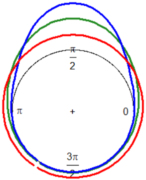
$$f(\theta;\mu,\kappa)=\frac{1}{2\pi I_0(\kappa)}\exp(\kappa\cos(\theta-\mu))$$




```
{circular}
mu<-circular(pi/2)
kappa<-2
curve.circular(dvonmises(x,mu,kappa),lwd=3,col="forestgreen")
kappa<-4
curve.circular(dvonmises(x,mu,kappa),add=T,lwd=3,col="blue")
```

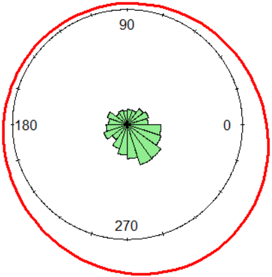
Densidade circular: von Mises

$$f(\theta;\mu,\kappa) = \frac{1}{2\pi I_0(\kappa)} \exp(\kappa\cos(\theta-\mu))$$



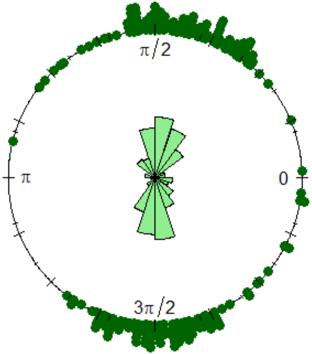

```
{circular}
mu<-circular(pi/2)
kappa<-2
curve.circular(dvonmises(x,mu,kappa),lwd=3,col="forestgreen")
kappa<-4
curve.circular(dvonmises(x,mu,kappa),add=T,lwd=3,col="blue")
kappa<-0.5
curve.circular(dvonmises(x,mu,kappa),add=T,lwd=3,col=2)
```

Densidade circular: von Mises




```
{circular}
mle.vonmises(pulgas,bias=TRUE)
```

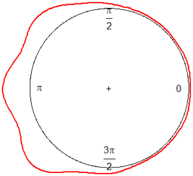
Cabalos do demo




```
{NPCirc}
data(dragonfly)
plot(circular(dragonfly),stack=T,col="darkgreen")
rose.diag(dragonfly,bins=20,add=T,col="lightgreen")
```

Estimación tipo núcleo da densidade

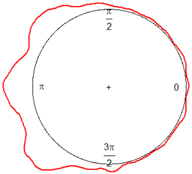
$$\hat{f}_\kappa(\theta) = \frac{1}{n}\sum_{i=1}^{n} K_\kappa(\theta - \Theta_i)$$




```
{NPCirc}
kern.den.circ(x,bw=50)
```

Estimación tipo núcleo da densidade

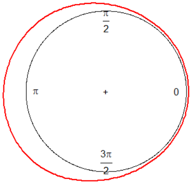
$$\hat{f}_\kappa(\theta) = \frac{1}{n}\sum_{i=1}^{n} K_\kappa(\theta - \Theta_i)$$




```
{NPCirc}
kern.den.circ(x,bw=200)
```

Estimación tipo núcleo da densidade

$$\hat{f}_\kappa(\theta) = \frac{1}{n}\sum_{i=1}^{n} K_\kappa(\theta - \Theta_i)$$




```
{NPCirc}
kern.den.circ(x,bw=5)
```

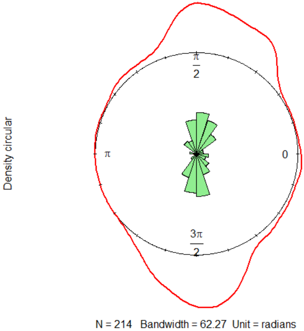
Selección do parámetro de suavizado

$$AMISE(\kappa) = \left\{ \frac{1}{16} \left[1 - \frac{I_2(\kappa)}{I_0(\kappa)} \right]^2 \int_0^{2\pi} [f^{''}(\theta)]^2 d\theta + \frac{I_0(2\kappa)}{2n\pi(I_o(\kappa))^2} \right\}$$

- Regra do polgar
- Regra plug-in
- Validación cruzada
- Bootstrap

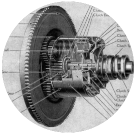
```
{NPCirc}
bw.rt(x)
bw.pi(x)
bw.CV(x)
bw.boot(x)
```

Estimación tipo núcleo da densidade




```
{NPCirc}
plugin<-bw.pi(dragonfly)
kern.den.circ(dragonfly,bw=plugin)
```

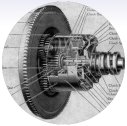
Regresión con variables circulares



- Explicativa Θ: ángulo de desequilibrio
- Resposta Y: peso corrector

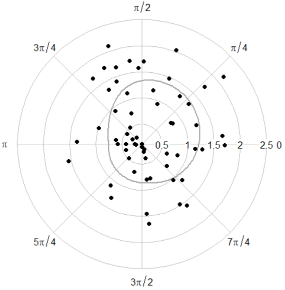
- Explicativa X: distancia
- Resposta Φ: dirección de migración

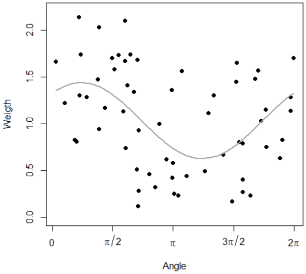
- Explicativa Θ: dirección do vento ás 6am
- Resposta Φ: dirección do vento ás 12am



Regresión circular-linear

$$Y_i = \beta_0 + \beta_1 \cos\Theta_i + \beta_2 \sin\Theta_i + \varepsilon_i$$

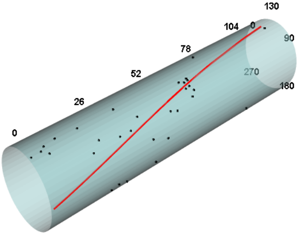





```
lm(peso~sin(angulo)+cos(angulo))
```

Regresión linear-circular

$$\Phi_i \sim vM(\mu_i, \kappa)$$
$$\mu_i = \gamma + 2\tan^{-1}(\beta X_i)$$



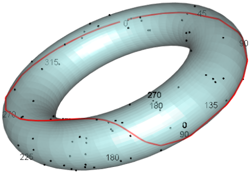
{circular}

```
lm.circular(type=‘‘c-l’’, y=cdirect, x=distance, init=0.0)
```

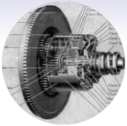
Regresión circular-circular

$$\cos(\Phi_j) = \gamma^c + \sum_{k=1}^{p} (\beta_k^c \cos(k\Theta_j) + \omega_k^c \sin(k\Theta_j)) + \varepsilon_j^c$$

$$\sin(\Phi_j) = \gamma^s + \sum_{k=1}^{p} (\beta_k^s \cos(k\Theta_j) + \omega_k^s \sin(k\Theta_j)) + \varepsilon_j^s$$




```
{circular}
lm.circular(type=‘‘c-c’’, y=noon, x=sixam, order=1)
```



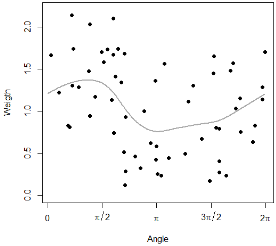
Regresión tipo núcleo circular-linear

- O modelo:

$$Y_j = m(\Theta_j) + \varepsilon_j$$

- Axuste trigonométrico **local**

$$\beta_0 + \beta_1 \sin(\Theta_j - \theta)$$

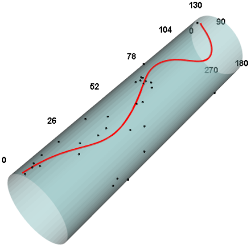


- Estimación:

$$(\hat{\beta}_0, \hat{\beta}_1) = \arg \min_{(a,b)} \sum_{j=1}^{n} K_\kappa(\theta - \Theta_j)[Y_j - (a + b\sin(\theta - \Theta_j))]^2$$

```
{NPCirc}
bwcv<-bw.reg.circ.lin(angulo,peso)
kern.reg.circ.lin(angulo,peso,bw=bwcv,method="LL")
```

Regresión tipo núcleo linear-circular



$$\Phi_i = [m(X_i) + \varepsilon_i](\text{mod } 2\pi)$$

- Estimador:

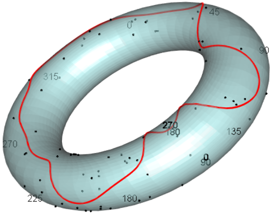
$$\hat{m}(x) = \text{atan2}(\hat{g}_1(x), \hat{g}_2(x)),$$

$$\hat{g}_1(x) = \frac{1}{n}\sum \sin(\Phi_i)W_h(X_i - x),\ \hat{g}_2(x) = \frac{1}{n}\sum \cos(\Phi_i)W_h(X_i - x)$$

```
{NPCirc}
bwcv<-bw.reg.lin.circ(distance,cdirect)
kern.reg.lin.circ(distance,cdirect,bw=bwcv,method="LL")
```

Regresión tipo núcleo linear-circular

$$\Phi_i = [m(\Theta_i) + \varepsilon_i] (\text{mod } 2\pi)$$



- Estimador:

$$\hat{m}(\theta) = \text{atan2}(\hat{g}_1(\theta), \hat{g}_2(\theta)),$$

$$\hat{g}_1(\theta) = \frac{1}{n} \sum \sin(\Phi_i) W_\kappa(\Theta_i - \theta), \; \hat{g}_2(\theta) = \frac{1}{n} \sum \cos(\Phi_i) W_\kappa(\Theta_i - \theta)$$

```
{NPCirc}
bwcv<-bw.reg.circ.circ(sixam,noon)
kern.reg.circ.circ(distance,cdirect,bw=bwcv,method="LL")
```

Paquetes de `R`

- `circular`: manexo de datos circulares
- `NPCirc`: técnicas non paramétricas

Bonus:

- `directional`: manexo de datos direccionais

Datos circulares con R

María Alonso Pena

Departamento de Estatística, Análise Matemática e Optimización

Universidade de Santiago de Compostela